State dependent choice

Article (Accepted Version)

Manzini, Paola and Mariotti, Marco (2015) State dependent choice. Social Choice and Welfare, 45 (2). pp. 239-268. ISSN 0176-1714

State Dependent Choice*

Paola Manzini†
University of St Andrews and IZA

Marco Mariotti‡
Queen Mary University of London

April 2015

Abstract

We propose a theory of choices that are influenced by the psychological state of the agent. The central hypothesis is that the psychological state controls the *urgency* of the attributes sought by the decision maker in the available alternatives. While state dependent choice is less restricted than rational choice, our model does have empirical content, expressed by simple 'revealed preference' type of constraints on observable choice data. We demonstrate the applicability of simple versions of the framework to economic contexts. We show in particular that it can explain widely researched anomalies in the labour supply of taxi drivers.

JEL codes: D01.

Keywords: Bounded rationality, procedural rationality, utility maximization, choice behavior.

*This paper radically revises and supersedes a previous paper entitled 'Moody Choice'. We are grateful for helpful comments to two anonymous referees, the Editor in charge Clemens Puppe, Attila Ambrus, Sophie Bade, Nick Baigent, Vince Crawford, Stephan Dickert, Yorgos Gerasimou, Andreas Gloeckner, Steffen Huck, Silvia Milano, Mauro Papi, Daniel Sgroi, Chris Tyson, Lin Zhang as well as several seminar audiences. We've had numerous inspiring discussions with Michael Mandler on related topics. Financial support through ESRC grant RES-000-22-3474 is gratefully acknowledged.

†School of Economics and Finance, University of St. Andrews, Castlecliffe, The Scores, St. Andrews KY16 9AL, Scotland, U.K. (e-mail: pm210@st-andrews.ac.uk).

‡School of Economics and Finance, Queen Mary University of London, Mile End Road, London E1 4NS, U.K. (e-mail: m.mariotti@qmul.ac.uk).

1 Introduction

In this paper we propose a theory of choices that are influenced by the psychological state in which the decision is taken. We wish to avoid as much as possible to commit to any specific psychological phenomenon. To do so, we imagine that an agent makes decisions by looking at the alternatives as carriers of attributes, or *properties*. This is fairly standard and is consistent with the Lancaster [36] tradition of consumer theory, with modern approaches to empirical demand analysis (e.g. Berry, Levinsohn and Pakes [6]) and with other abstract modern theories of choice (e.g. Dietrich and List [15]).

What gives leverage to our theory is the assumption that properties are looked at sequentially by the agent. For example, when searching for a restaurant, the agent may first check only the restaurants that are nearby, then focus on the French ones among these, then see whether there is any nearby French with acceptable prices, and so on. In general, following a list of desirable properties, the agent progressively eliminates at each stage alternatives that do not possess the relevant property, until a set of acceptable alternatives is found.[1] This is a case of 'sequentially rationalisable choice' (Manzini and Mariotti [41]) and we showed in Mandler, Manzini and Mariotti [40] (henceforth MMM) that the decisions resulting from this type of procedure are indistinguishable from those of a standard utility maximising agent. In other words, MMM cannot explain any violation of the Weak Axiom of Revealed Preference (WARP). Thus, while it has the methodological value of providing a procedural foundation for utility maximisation, the MMM model does not go beyond the standard model in terms of the phenomena it can explain. In this paper, we extend the explanatory power of the MMM model by dropping the assumption that the order in which the properties are applied is fixed. So, to continue with the restaurant example, we allow the agent to consider as the most urgent property sometimes price, sometimes the type of cuisine, and sometimes the location, as well as to have sometimes more drive for one type of cuisine and sometimes for others. Or, we allow an investor to sometimes look first at the safety features of an investment and some other times (perhaps in an irrationally exuberant mood) to consider the maximum return as the most important property. In this way, while obtaining a theory that is not vacuous, we are able to capture interesting types of 'non-standard' behaviour that are determined by a psychological state,

[1] See Mandler, Manzini and Mariotti [40] for a discussion of the psychological foundations of this procedure.

modelled as the relative urgency of the properties sought by the agent. Furthermore, the theory becomes sufficiently flexible to be applied to economics settings in a fruitful way, explaining some observed anomalies that are difficult to accommodate in the standard utility maximisation model.

That the psychological state of the agent has an effect on the cognitive process underlying choice is intuitive and is an increasingly accepted feature of economic models. For example, the relationship between 'mood' and choice is documented in an impressively large body of empirical evidence from psychology and economics.[2] In the large literature on menu choice (Kreps [34]; Dekel, Lipman and Rustichini [14]), an agent anticipates being in one of different preference determining states. And there exist several models of the effects on choice of specific psychological states, such as 'anticipatory feelings' (Caplin and Leahy [8]). In the MMM model the cognitive process underlying choice is summarised by the order in which properties are applied. Therefore, if it is granted that psychological states do matter, it is natural to take the order of the relevant properties as a variable rather than a fixed element of the model.

One of our main aims is to arrive at a general, flexible framework to incorporate the effects of the disparate psychological factors affecting choice, with a view to teasing out the empirical restrictions entailed by the dependence on these factors *per se* (that is, independently of the specificity of the factor). One 'natural' approach might appear that of using a state dependent utility function. However, one difficulty of this approach is that it has an empty empirical content when states depend on choice menus (see section 6). For this reason, while still avoiding to commit to a specific interpretation of what a psychological state is, we use the richer primitives of the MMM model rather than the behaviourally equivalent utility maximisation model. By doing so, we obtain a framework in which behaviour always presents systematic patterns even when subject to psychological influences. Such patterns can be identified by means of direct observation of choice data, through tests that are simple variations of standard tests in revealed preference theory.

We call *mindset* the set of properties considered in general by the agent, 'the things that make the agent tick'. A glutton would have to change his mindset/personality, not just the contingent factors affecting choice, to start paying attention to the fat content of his diet. A foolhardy agent will always ignore the safety aspects of his options. While the mindset is fixed throughout all choices, how the relevant properties are ranked

[2] A sample of publications is: Capra [10], Erber et al. [19], Ifcher and Shaghaghi [26] Isen [27], Kahn and Isen [28], Kirchsteiger et al. [30], Mayer et al. [38], Mittal and Ross [47], Nygren [48], Nygren et al. [49], Oswald et al. [50], Thayer [54], [55], Williams and Voon [56]

with respect to one another can vary. We call the order in which properties are considered a *state*. This methodology allows us to describe an agent with a recognisable identity (in the form of a fixed set of values) who makes different choices in different conditions: an alternative with the property that is top in a state might be selected in that state, but discarded in a state that prioritises different properties that this alternative does not possess.

To give a flavour of how the model works, the investor depicted in Figure 1 can be in two states, bearish or bullish. His mindset consists of just two 'threshold' properties: minimum acceptable return and maximum acceptable risk. In a bearish state, the investor puts risk limitation at the forefront and selects portfolio b, which is the only one of the two to have the relevant property, while in a bullish state she looks first for a minimum return, selecting portfolio a.

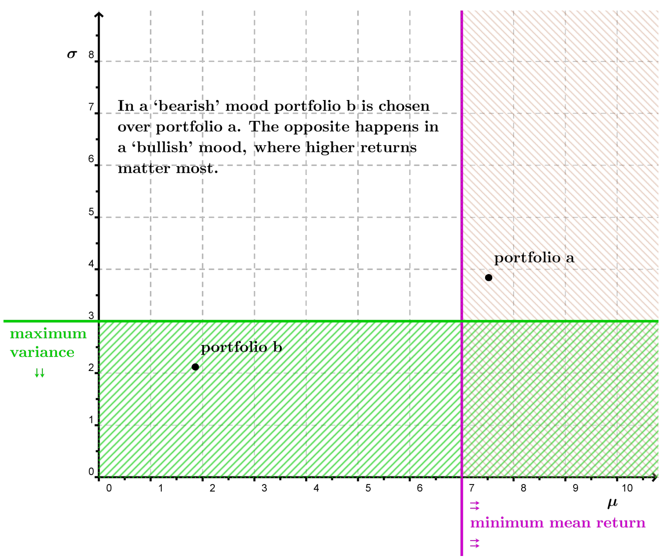

Figure 1: Bulls and bears

This cognitive process underlying choice extends to situations with more numerous and more complex properties.

How are states determined? We consider two models that capture two distinct plausible possibilities. In the first model, a state is *menu driven*, that is, triggered by the menu under consideration. This is a most natural situation in consumer choice, as sellers may and do manipulate menus to their advantage (e.g. using "asymmetric dominance" effects), but it may occur in other contexts too. For example, if an agent finds it costly to contemplate a menu before selecting an alternative from it (e.g. Ergin and Sarver [20]), menus with different contemplation costs may induce different states.

In the second model, a state is *environment driven*, that is, determined by factors exogenous to the choice menu, and the agent is observed to make multiple choices from the same menu in different states. In figure 1 imagine that a climate of 'irrational exuberance' (a memorable instance of a collective psychological state) triggers an appetite for risk and hence makes the 'high return property' the more urgent one. The environmental state model assumes that, as in most cases outside controlled experiments, we cannot observe the state of an agent making a choice.

We show that the model in which states are menu driven is equivalent to choice data satisfying a single property (Togetherness). An interesting aspect of this property is that it is intermediate in strength between two textbook properties: it is weaker than the Weak Axiom of Revealed Preference (WARP) and stronger than Sen's property β (section 3). Concretely, think of a competitive consumer choosing from budgets. Togetherness says that all those bundles that are acceptable at some budget either remain (if still available) all acceptable or become all unacceptable following a price/income change. Togetherness implies in particular that an agent who reveals himself willing to engage in a *sequence* of trades, say leading from x to y, is also willing to engage in a *direct* trade between x and y. So in particular a welfare planner can still use such 'revealed indifferences' as a guidance, as he would with a standard agent.

The second model, where states are environment driven, admits multiple observations of choice from the same menu, made in different states. We show that the set of possible choice observations (in all possible states) fails Independence of Irrelevant Alternatives and other consistency properties, yet is again fully characterised by a single property (state dependent WARP, or sd-WARP) that is intermediate between two standard properties: it is weaker than WARP and stronger than Sen's Property α. Once again, consider a competitive consumer. sd-WARP states in that context that if a bundle is always (i.e. in any state) rejected at some prices and income (p, M), then it is always rejected at any other budget for which *all* the bundles demanded (in some state) at (p, M) are affordable (section 4).

An advantage of the axiomatic approach is that it automatically generates a recipe for direct empirical tests of the theory. Secondly, it permits simple and sharp behavioural distinctions between our theory and other axiomatic theories of psychologically driven choice, even using a very limited range of menus and observations. In this vein, we show how to distinguish between our model and a model of *indecisive* behaviour (section 6.3).

Though abstract, our framework is workable and can yield novel economic insights. In section 5 we specialise the models to use them in a specific economic settings. We look at a long-standing debate on labour supply responses by workers who (like taxi-drivers) can choose their supply in a non-lumpy way.[3] Even a very stripped down

[3]Some of the relevant literature is cited in section 5.

version of the model (that postulates 'target-driven' behaviour as suggested in the empirical literature) can accommodate the 'anomaly' of negative wage elasticities. The new key insight we provide lies in the asymmetry between income targets, which can be 'unrealistic', and leisure targets, which have a physical bound (there are so many hours in a day). This interplay means in the model that workers can display a whole range of wage elasticities depending on the magnitude of wage changes and independently of any role of expectations (which conversely play a key role in reference-dependent explanations). This application also yields two novel empirically testable predictions. Similar simplifications of the theory can be easily applied to saving decisions and consumer theory.

2 Mindset and states

Fix a finite set of alternatives X. Given the collection Σ of all nonempty subsets of X, a choice function is a correspondence c that associates with each $A \in \Sigma$ (a *menu*) a nonempty set $c(A) \subseteq A$, the agent's observed selection from A. To build a model of state dependent choice, we begin by assuming, as in Mandler, Manzini and Mariotti [40] (MMM), that the agent makes choices by sequentially going through a checklist of 'properties' of alternatives (properties are intended as synonymous with 'attributes'). At each step, he discards the alternatives that lack the relevant property. In spite of it being procedural (it is a particular case of procedures such as sequential rationalisability as in Manzini and Mariotti [41] and those in Apesteguia and Ballester [4]), this decision model is shown in MMM to be equivalent to ordinary preference maximisation: an agent has a checklist if and only if he maximises a preference relation. Any checklist corresponds to a preference, and vice versa. However the checklist model has richer primitives than ordinary preference maximisation. This feature permits to distinguish, unlike a preference, between the more stable traits of the agent's personality (mindset), and the more variable aspects (states).

We identify a property with the set of alternatives that possess that property. So formally a property is a subset $P \subseteq X$, and we say that x has property P whenever $x \in P$. E.g. the property 'sweet' consists of all the objects in X which are sweet. A **mindset** $\Gamma \subseteq 2^X \backslash \varnothing$ is a set of mutually distinct properties. A mindset Γ is **nested** if for all $P, Q \in \Gamma$ we have either $P \subseteq Q$ or $Q \subseteq P$. Given a mindset Γ, a **state** is a strict linear order $<$ of Γ.

In the revealed preference tradition, we assume throughout that the state itself in which a choice is made is not necessarily observable: only the resulting choice is.

3 The state is determined by the menu

In our first model the state is triggered by the menu. Multiple choices from a menu are allowed, and they are interpreted as being made always in the same state (the one triggered by the menu itself). Given a mindset Γ and a menu $A \in \Sigma$, denote by $<_A$ the state induced by A (which we refer to as 'a state for A'). Denote by P_0^A the first property in state $<_A$ (the $<_A$ $-$least element of Γ).[4]

A **menu checklist** is a pair $\left(\Gamma, \{<_A\}_{A\in\Sigma}\right)$ of a mindset and a collection of states, one for each menu. Given $A \in \Sigma$ and a menu checklist $\left(\Gamma, \{<_A\}_{A\in\Sigma}\right)$, we can define inductively a series of 'survivor sets':

$$S_A\left(P_0^A, <_A\right) = A$$

and

$$\text{For all } P \in \Gamma \backslash P_0^A\text{: } S_A\left(P, <_A\right) = \begin{cases} \bigcap\limits_{Q <_A P} S_A\left(Q, <_A\right) \cap P & \text{if } \bigcap\limits_{Q <_A P} S_A\left(Q, <_A\right) \cap P \neq \varnothing \\ \bigcap\limits_{Q <_A P} S_A\left(Q, <_A\right) & \text{otherwise} \end{cases}$$

That is, when facing the menu A, the agent's state for that menu identifies the order of the properties in the checklist. The agent scans the checklist, and when considering each property he only retains the alternatives which have that property, if any. Otherwise he retains all the alternatives that have survived until that stage. Then the agent moves to next stage. When convenient we omit denoting the linear order in the sets $S_A\left(P, <_A\right)$ and simply write $S_A\left(P\right)$.

Definition 1 *A choice function c on Σ is a* ***menu driven state dependent choice*** *if there exists a menu checklist $\left(\Gamma, \{<_A\}_{A\in\Sigma}\right)$ such that for all $A \in \Sigma$, for some property $P \in \Gamma$,*

$$\begin{aligned} S_A(P) &= S_A(Q) \textit{ for all } Q \in \Gamma \textit{ with } P <_A Q \\ c(A) &= S_A(P) \end{aligned} \tag{1}$$

[4]While in this paper we confine ourselves to finite sets, the definitions are written so as to immediately extend to the infinite cases.

In words, a menu driven state dependent choice is such that all those alternatives that are chosen from a menu are in the 'last' survival set constructed on the basis of the relevant state for that menu.

While each menu triggers a different order in which the various properties are considered, MMM consider checklists that are independent of the menu. A choice function c on Σ **has a checklist** $(\Gamma, <)$ if it is a menu driven state dependent choice with menu checklist $\left(\Gamma, \{<_A\}_{A\in\Sigma}\right)$ such that $<_A = <_B = <$ for all $A, B \in \Sigma$. It turns out that a checklist expresses the standard notion of rationality in economics, preference maximisation. A choice function c maximises a preference if there exists a weak order[5] $\succsim$ on X such that, for all $A \in \Sigma$, $c(A) = \{x : x \succsim y \text{ for all } y \in A\}$:

Proposition 1 *(MMM, [40]): A choice function c has a checklist if and only if it maximises a preference.*[6]

A choice function may be a menu driven state dependent choice even if it has no checklist (see appendix 1). In light of the above proposition, this means that the model we propose can explain behaviour that is not preference maximising. The following example illustrates:

Example 1 *A consumer enters an 'exuberant' state when he faces large menus or menus composed entirely of luxury items, but is in a thrifty state when a thrifty item is available in a small menu. As a consequence he will for instance choose an expensive food item from a hefty restaurant list, and a modest entree from a shorter one. Formally, let $X = \{x, y, z\}$, where x and y are luxury items, z is thrifty, and only the grand menu is large. So*

$$c(X) = c(\{x,y\}) = \{x,y\} \text{ and } c(\{x,z\}) = c(\{y,z\}) = \{z\}$$

The choice function c cannot maximise any weak order ($c(X) = \{x,y\}$ would imply $x \succsim z$ while $c(\{x,z\}) = \{z\}$ would imply $z \succ x$) and therefore by Proposition 1 it cannot have a checklist. Nevertheless it is a menu driven state dependent choice with the menu checklist $(\{\{x,y\}, \{z\}\}, <_A)$ and states $\{x,y\} <_A \{z\}$ for $A \in \{X, \{x,y\}\}$ and $\{z\} <_A \{x,y\}$ for $A \in \{\{x,z\}, \{y,z\}\}$.

We also make at this stage an observation that may appear surprising: in spite of the fact that each checklist corresponds to a preference, it is *not* necessarily the case that

[5] A weak order is a a complete transitive relation.

[6] This result is proved in MMM in greater generality than for the domain considered here. The connections with the classical theory of lexicographic preferences are explained in that paper.

an agent who maximises a menu dependent preference is captured by a menu driven state dependent choice. This implies that the model does have empirical restrictions (unlike the model of menu dependent preferences). Details are in section 6.

3.1 Menu driven state dependence may not manifest itself in behaviour

State variations are not necessarily expressed in observable choice behaviour. In particular, a necessary condition for state changes to be observable is that the properties cannot be ranked as more or less permissive, i.e., they are *not nested*.

Example 2 *There are two properties you look at when selecting from a restaurant menu: 'recommended by a friend' and cheapness (all dishes are sufficiently appealing). The rice dish (r) has been recommended by a friend, while steak (s) and tacos (t) have not been recommended. Rice and tacos are cheap and steak is not. In a trusting state (the recommendation first springs to mind) rice is selected over both tacos and steak, and tacos are selected over fish. But evidently in a conservative state (cheapness first), you would make exactly the same choices! (Rice and tacos survive the first cheapness test and then rice is selected on the basis of the recommendation). Formally,*

$$c(\{r,s,t\}) = c(\{s,r\}) = c(\{r,t\}) = r \text{ and } c(\{s,t\}) = t$$

and the data are explained by the mindset $\{\{r\},\{r,t\}\}$ *both in the state* $\{r\} <_A \{r,t\}$ *for all* A *and in the 'opposite' state* $\{r,t\} <_A \{r\}$ *for all* A*.*

This fact holds in general: when the mindset is nested, the state does not affect choice, an agent affected by states behaves exactly like one that is not (but has the same mindset).

Proposition 2 *Let c be a menu driven state dependent choice with nested mindset. Then c maximises a preference.*

(The proofs of all propositions and claims, trivial and less trivial, are in the appendix). A leading example of nested properties is provided by a textbook utility maximiser who uses as properties the upper contour sets of the utility function. This case has a natural procedural interpretation: the agent is in fact a satisficer who at each stage s sets a threshold numerical satisfaction target t_s. At stage s, the agent keeps only the satisficing alternatives (those that meet the target t_s), if any, and otherwise he keeps all of them. In the next stage he revises the satisfaction threshold. There is no need to specify the revision rule, precisely because the properties are nested. The best alternatives in a menu will never be eliminated: if, when considering a property P, the set of survivors from the previous stages contains some alternatives that are in P, then the best alternatives must be among them. The state, in this interpretation, manifests itself in the initial property t_1 and the revision rule adopted: sometimes the agent will start with ambitious targets, and sometimes with more modest ones; sometimes he will react to the lack of satisfactory alternatives by radically revising down the target, and sometimes he will hold firm. This agent is procedural as well as subject to state changes but his choice behaviour never appears to an external observer to be swayed by this instability!

But the invariance of choice to state changes does not necessarily rest on nested states. In specific economic settings there are natural non-nested properties that - at least within a parameter range - generate choice invariance.

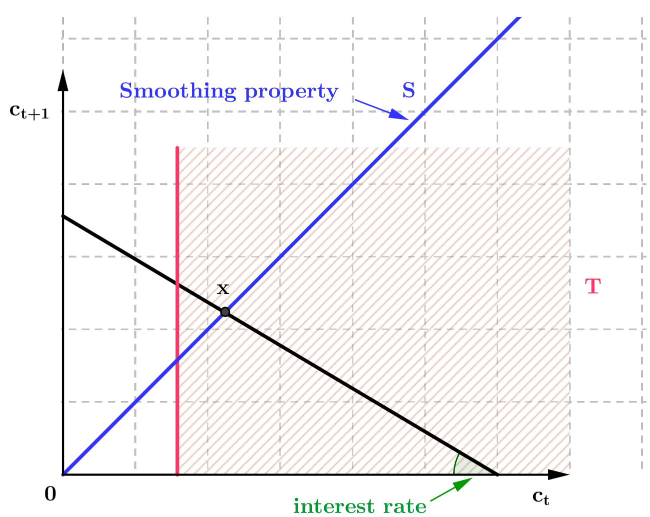

Figure 2: State dependent intertemporal consumption

In Figure 2 a saver in the standard consumption today and tomorrow (c_t, c_{t+1}) space considers first two properties: intertemporal consumption smoothing (S, the 45^0 line) and satisfaction of immediate urges (all consumption patterns to the right of the vertical line, the set T). For the interest rate indicated (and all those sufficiently close to it), choice is invariant to whether S is applied before T or viceversa, and is always the point x.

For much lower interests rates, however, the state would become relevant. The dependence of the relevance of states on economic parameters in examples of this kind can be exploited to derive interesting economic implications and explain some anomalies. We elaborate on this point in section 5.

3.2 The observable behavioural implications of menu driven state dependence

Once we step aside of the special cases studied in the previous section, however, menu driven state dependence does effect choices that cannot be rationalised in a standard manner.

Example 3 *Rice (r) and tacos (t) are both choosable when steak (s) is also available, but rice alone is chosen over tacos when these are the only two available dishes. Formally*

$$c(\{r,s,t\}) = \{r,t\} \text{ and } c(\{r,t\}) = \{r\}$$

The choice from $\{r,t\}$ implies that, if there existed a menu checklist for c, then its mindset would contain a property P which 'separates' r and t, i.e. $r \in P$ and $t \notin P$. But then, whatever the state, such a property would sooner or later also separate r and t in $\{r,s,t\}$, contradicting $r,t \in c(\{r,s,t\})$.

The reasoning in the above example suggests a necessary property:

Togetherness: If an alternative x is rejected from a menu from which another alternative y is chosen, then x cannot be chosen when y is chosen. Formally, for all $A, B \in \Sigma$: $[x \in A \backslash c(A), y \in c(A), y \in c(B)] \Rightarrow x \notin c(B)$.

This property is equivalent to saying that if x and y are both chosen from a menu, then from any other menu x and y are either both chosen or both rejected, hence the name Togetherness. It is intermediate in strength between two classical revealed preference properties: the Weak Axiom of Revealed Preference (WARP) and property β. These are defined as follows:

WARP: If an alternative x is rejected from a menu from which another alternative y is chosen, then x cannot be chosen when y is available. Formally: For all $A, B \in \Sigma$: $[x \in A \backslash c(A), y \in c(A), y \in B] \Rightarrow x \notin c(B)$.

Property β: If an alternative x is rejected from a menu from which another alternative y is chosen, then x cannot be chosen from a smaller menu from which y is chosen. Formally, for all $A, B \in \Sigma$: $[B \subset A, x \in A \backslash c(A), y \in c(A), y \in c(B)] \Rightarrow x \notin c(B)$.

These axioms share the same conclusion as Togetherness, but Togetherness reaches it from a weaker (resp. stronger) premise than Property β (resp. WARP).

While apparently weak, Togetherness is fairly restrictive. For example, beside being stronger than Property β as noted, it is also (strictly) stronger than a natural variant of WARP introduced by Ehlers and Sprumont [17], which requires that if the agent chooses x and rejects y from a menu, then he cannot choose y and reject x from another menu.[7]

It turns out that Togetherness fully characterises our model.

Proposition 3 *A choice function is a menu driven state dependent choice if and only if it satisfies Togetherness.*

This result clarifies the sense in which our model yields testable conditions for state dependent behaviour. An agent, no matter how state dependent, will never be observed to accept two alternatives from one menu while rejecting only one of the two from another menu. And conversely, every time that this pattern is not observed we can explain observed choices by means of the model. Suppose I'm willing to go to the theatre or to the cinema when the only other alternatives are to work or to keep the appointment with the dentist, but choose to keep the appointment with the dentist when the other alternatives are cinema, theatre, work, and a visit to a friend in the hospital. In this example we have a violation of WARP but not of Togetherness: the switch from the choice of cinema or theatre to the choice of seeing the dentist can be imputed to a shift in state. It could be, for instance, that the presence of a friend in the hospital, while not providing me with a sufficiently strong reason to visit her, puts me in a pensive mood, focussing my mind away from entertainment choices.[8]

Any single-valued choice function satisfies Togetherness, so Proposition 3 does not impose testable restrictions on this type of data. However, the result applies to the most abstract version of the theory, with no restriction on how states may depend on menus. It is not difficult to build specialisations of the theory which are restrictive even for single-valued choice data. A first example comes from Proposition 2: if properties are nested, the theory can be falsified by single-valued choice data, through violations of WARP. A second quite natural example is a generalisation of the well-known 'Luce and Raiffa's dinner', whereby the presence of frog's legs on a menu triggers a change in preference. Suppose that a menu may or may not contain a 'trigger'

[7] The easy proof of this assertion is left as an exercise for the reader.

[8] As it often happens, choice data alone cannot be the ultimate arbiter in the selection of the model. For example, 'rationalisation' in the sense of Cherepanov, Feddersen and Sandroni [11] could be the 'true' explanation of the behaviour just described. Context and non-choice data will help model selection.

alternative that alters the psychological state (beside the Luce and Raiffa dinner example, the presence of an expensive bottle of wine may make a diner more inclined to splash out, or noticing a high paying job advertisement may put a job applicant in the mood for more ambitious applications). Given an alternative $t \in X$ define a **trigger checklist** as a menu checklist $(\Gamma, \{<_t, <_o\})$ with two states, with state $<_t$ applying to all menus A such that $t \in A$, and the 'ordinary' state $<_o$ applying whenever the decision maker is choosing from a menu which does not contain the trigger alternative (e.g. we can see $<_o$ as a 'neutral' state and $<_t$ as an 'excited' state). Call a menu driven state dependent choice with a trigger checklist a **trigger state dependent choice**. There are single-valued choice functions which are not trigger state dependent choice. Consider for instance the choice function c defined on $X = \{x, y, w, z\}$ as $c(X) = w$; $c(xyw) = x = c(xwz)$; $c(xyz) = y = c(ywz)$; $c(yz) = z$. Recall that each checklist determines a preference order. Whichever order is applied to X must rank w above any other alternative; consequently a different order must be applied to all three-sets. However $c(xyw) = x = c(xwz)$ implies that x ranks above all other alternatives, while $c(xyz) = y = c(ywz)$ implies that y ranks above all other alternatives, a contradiction.

In general, it easy to prove[9] that a choice function is a trigger state dependent choice if and only if it there exists an alternative that partitions the domain in two subsets such that WARP applies to each subdomain.[10]

The relation $\approx$ introduced in the proof of Proposition 3 can be interpreted as 'behavioural indifference': $x \approx y$ requires x and y to be never separated by choice. A menu driven state dependent agent satisfies the transitivity of the behavioural indifference $\approx$. So in particular he is willing to carry out in one step a trade, such as between x and y, when he has implicitly (via $\approx$) revealed his willingness to carry out a sequence of pairwise trades leading from x to y. Moreover he will be willing to carry out explicitly the implicit sequence of trades in any order. This is because the implicit trades which are acceptable to a menu driven state dependent agent are all and only the trades between alternatives that have exactly the same properties, and 'having the same properties' is a symmetric and transitive relation.[11] Of course, alternatives can share the same rele-

[9]The proof is available from the authors.

[10]Formally, c must satisfy the following weakening of WARP, where for any alternative $x \in X$ we define $\Sigma_x = \{S \subseteq X : x \in S\}$ and $\Sigma_{\neg x} = \Sigma \backslash \Sigma_x$:

t-WARP (trigger WARP): There exists $t \in X$ such that for $i = x, \neg x$: For all $A, B \in \Sigma_i$: $[x \in A \backslash c(A), y \in c(A), y \in B] \Rightarrow x \notin c(B)$.

[11]Mandler [39] shows how indifference can be distinguished from incompleteness by observing the trades an agent is willing to carry out.

vant properties and be physically different: for example, walking, cycling and taking a bus can all belong to a 'cheap leisurely means of transport' behavioural indifference class for an agent.

These observations are important for welfare analysis. Like in the standard case, a planner faced with the task of choosing between implementing x or y may let himself be guided by the sequences of behavioural indifferences of a possibly menu driven state dependent agent without fear of hurting the agent's welfare. Unlike standard welfare analysis, the planner can no longer always use as a guide for welfare judgements the behavioural strict preferences if he suspects that the agent is menu driven state dependent. Nevertheless, some choices may provide information about strict welfare judgements. In the previous example, work is *never* chosen, whereas all other three alternatives which are available in different states (and we know that the state has changed across menus since WARP has been violated) are chosen under at least one state. We deduce that, whatever the hierarchy of importance among the properties sought in the alternatives, work never offers a crucial property that some other alternative does not offer, and there are crucial properties that other alternatives offer but work does not. Work appears thus a good candidate to be declared welfare inferior to the other three alternatives available in both states. We cannot make the same inference regarding visiting a friend in the hospital, since such a choice was not available in the first state. This type of reasoning is a variant of Bernheim and Rangel's [5] approach to 'behavioural welfare economics' (see Manzini and Mariotti [43] for a discussion on welfare and bounded rationality. See also Masatlioglu et al. [46] and Rubinstein and Salant [52]).

4 The state is determined by the environment

In our second model the state does not depend on the menu, but we allow several observations of choice from an A, each time in a different (unidentifiable) state.[12] A mindset is as before a set Γ of properties, and a state is a linear order $<_m$ of Γ. We allow for the possibility that each observation is itself multivalued.

Specifically, let M be the set of states in which any menu A is considered. A pair $\left(\Gamma, \{<_m\}_{m\in M}\right)$ is called an **environmental checklist**. Let $\gamma\left(A, <_m\right)$ denote the choice from A in state $<_m \in M$. Analogously to before, $\gamma\left(A, <_m\right)$ is determined by a sequence

[12]As in the previous section, our setup is static. Laibson [35] studies a dynamic model of what we would call an 'environment driven state' triggered by binary cues.

of successive eliminations. Denote by P_0^m the first property in state $<_m$. The survivor sets are defined by

$$S_A\left(P_0^m, <_m\right) = A$$

and

$$\text{For all } P \in \Gamma \backslash P_0^m\text{: } S_A\left(P, <_m\right) = \begin{cases} \bigcap\limits_{Q <_m P} S_A\left(Q, <_m\right) \cap P & \text{if } \bigcap\limits_{Q <_A P} S_A\left(Q, <_m\right) \cap P \neq \varnothing \\ \bigcap\limits_{Q <_m P} S_A\left(Q, <_m\right) & \text{otherwise} \end{cases}$$

Then, for all $A \in \Sigma$, $\gamma(A, <_m)$ is defined as follows:

$$\begin{aligned} \gamma(A, <_m) &= S_A(P, <_m), \text{ where } P \in \Gamma \text{ is such that} \\ S_A(P, <_m) &= S_A(Q, <_m) \text{ for all } Q \in \Gamma \text{ with } P <_m Q \end{aligned} \tag{2}$$

We begin by noting that the functions $\gamma\left(., <_m\right)$, which describe the observations conditional on one state, are consistent in the sense that they satisfy IIA:

IIA: If some alternatives of a larger menu are still available in a smaller menu, the alternatives chosen from the smaller menu are the available ones which are chosen from the larger menu. Formally, for all $A, B \in \Sigma$: $\left[A \subset B, c\left(B\right) \cap A \neq \varnothing\right] \Rightarrow c\left(A\right) = c\left(B\right) \cap A$.

Then:

Proposition 4 *For all $<_m \in M$, $\gamma\left(., <_m\right)$ satisfies IIA.*

This result is implied by Proposition 1 and standard properties of preference maximisation, but we also give a direct proof in the Appendix.[13] Unfortunately, Proposition 4 is often not of practical use: although each $\gamma\left(., <_m\right)$ satisfies IIA, the specific state in which choice was made is most likely unobservable, as an external observer can often be expected to have only choice data, not state data, available. Thus, the object $c\left(A\right)$ is now interpreted as the collection of all the choices the agent makes from A in all possible unobservable states:

[13] While we are working for simplicity in the full domain, so that IIA and WARP are equivalent, the result is independent of this domain restriction.

Definition 2 *A choice function c is an* ***environment driven state dependent choice*** *if there exists an environmental checklist* $\left(\Gamma, \{<_m\}_{m \in M}\right)$ *such that for all* $A \in \Sigma$

$$c(A) = \bigcup_{<_m \in M} \gamma(A, <_m) \text{ for all } A \in \Sigma$$

Abstracting from the fact that the choices γ are generated here by a checklist and not by a utility, this framework parallels Salant and Rubinstein's [53] 'choice with frames' (or Bernheim and Rangel [5] similar framework of 'choice with ancillary conditions'), where a choice correspondence is interpreted as including, for each A, all the single-valued choices made from A in some frame. The definition here is similar, but we allow the choice in a state, $\gamma(A, <_m)$, to be multi-valued.[14] This means that, at a substantive level, the models differ, as there may not be a one-to-one correspondence between the frames and the states that explain a given set of observations. Details are in section 6.

4.1 Environment driven state dependence causes behaviour inconsistency

We search for standard 'consistency' properties that c may satisfy. First of all, it is easy to show that c does not inherit IIA from the $\gamma(., <_m)$. Alternatives which are not chosen, in any state, from a larger menu, may be chosen, in some state, from a smaller menu in which they are available. We illustrate this with an example which shows, more specifically, that c fails to satisfy two classical basic consistency properties implied by IIA. One is Property β already defined, and the other is:

Expansion: An alternative chosen from two menus must still be chosen when the two menus are merged. Formally, for all $A, B \in \Sigma$: $c(A) \cap c(B) \subseteq c(A \cup B)$.

Suppose there are three properties you look at when selecting from a restaurant menu: recommended by a friend, cheapness, and perceived appeal. Rice and steak have been recommended, rice and tacos are cheap, and steak and tacos are appealing. In a trusting state you sift through the properties in this order (first recommendation, then cheapness, then appeal). In a confident state you switch the order of the first and last property: you prefer to rely on your own judgement and the last thing you look at

[14] Below (section 6) we explore the relationship with Salant and Rubinstein's [53] in more detail. Bernheim and Rangel [5] also allow choice to depend on information beyond the feasible set, although their focus is not on the properties of c but rather on the welfare inferences that could be made by observing c.

is friends' recommendations.

So in a trusting state the steak, which has been recommended by a friend, is selected over the tacos, which have not been recommended. And the steak is also selected, in a confident state, over the rice dish, because the latter is less appealing.

However both tacos and rice are cheaper than steak. When all three dishes are on the menu, you are never observed to select steak (in violation of Expansion and Property β). The reason is that when you are trusting, steak and rice, which have both been recommended, survive the first elimination round and then rice is selected on the grounds of cheapness. And when you are confident, steak and tacos, which are both appealing, are shortlisted, to finally select the cheap tacos. Formally:

Claim 1 *c does not satisfy Expansion, nor Property β.*

Obviously, since Togetherness is stronger than Property β, this result also shows that c fails Togetherness.

4.2 Consistency of environment driven state dependent choice

We now show that environment driven state dependent behaviour is nonetheless subject to strong empirical restrictions. Of course if the mindset comprised only nested properties, we would have a result analogous to Proposition 2.

Proposition 5 *Let c be an environment driven state dependent choice with a nested mindset. Then c maximises a weak order $\succsim$ on X.*

Property α: All the alternatives chosen from a larger menu are still chosen when available in a smaller menu. Formally, for all $A, B \in \Sigma$: $[A \subset B, c(B) \cap A \neq \varnothing] \Rightarrow c(B) \cap A \subseteq c(A)$.

As we shall see, if c is an environment driven state dependent choice, it must satisfy Property α. Intuitively, if in some state you pick steamed salmon from a menu, it means that in that state steamed salmon fulfils some crucial property which the other alternatives do not fulfil, and this will continue to be the case even in subsets of that menu. Yet property α is not sufficient to characterise c (see appendix B.1). We need a significantly stronger condition that must be fulfilled by an environment driven state dependent choice:

sd-WARP: If an alternative x is rejected from a menu A, then x is rejected from any other menu that contains *all* the alternatives chosen in A. Formally, for all $A, B \in \Sigma$: $[x \in A \backslash c(A), c(A) \subseteq B] \Rightarrow x \notin c(B)$.

For example, if you were observed to choose sometimes seabass and sometimes a vegetarian meal, but never salmon, from a menu, then you will not choose salmon from any new menu that includes both seabass and the vegetarian meal. If your behaviour is determined by a state, it is easy to understand why this must be the case. Whatever state you are in, your choices from the old menu reveal that the first property that discerns between salmon and seabass (resp., the vegetarian meal) is such that seabass (resp., the vegetarian meal) has it while salmon lacks it.

WARP strengthens sd-WARP simply by replacing the entire choice set $c(A)$ with *any* alternative contained in it. For a fully rational agent any chosen element is representative of the class of chosen elements, but not so for an environment driven state dependent agent. Note also that sd-WARP implies Property α (see appendix B.2). The main result of this section is:

Proposition 6 *A choice function c is an environment driven state dependent choice if and only if it satisfies sd-WARP.*

Changes in mental state thus are compatible with choice exhibiting a significant degree of consistency. A violation of sd-WARP informs us that the agent's choices cannot be explained by 'preference maximisation plus states'. And, sd-WARP exhausts the testable implications of environmentally induced changes in mental state. Of course, as sd-WARP and Togetherness are logically independent conditions, so are the two models (see appendix B.3).[15]

5 A short application

Our treatment has been rather abstract so far. How it could be applied to the analysis of specific problems may be still unclear. Yet the same objection could be raised by a beginning economics student who is taught that the standard rational consumer maximises preferences: the concept of preference maximisation, while testable even in its general formulation, acquires concrete meaning only when it is applied, e.g. once a

[15]In appendix D we also establish, as an homage to aficionados of choice theory, a connection between our result and an old result from the Russian school.

quasi-concave utility function is maximised over a budget set. Our framework, while also testable in its general formulation, offers a flexibility similar to standard preference maximisation (to which state independent choice collapses): in the same way as e.g. Cobb-Douglas preferences rationalise a rigid labour supply, specific mechanisms of generating states that are appropriate for a given economic environment can help understand the logic of our theory, how it works, and the qualitative predictions and novel insights it can provide in the chosen settings. We pursue this point in the exercise below, which uses primitives similar to those of other theories (target variables). The different cognitive mechanism postulated by our theory accounts for the different implications.

5.1 NYC cab drivers' supply with target income and leisure

We consider the way in which unconstrained suppliers of labour (notably, taxi-drivers), as opposed to agents who necessarily have 'lumpy' supplies (e.g. factory workers), react to changes in wage. This is a long-standing, theoretically and empirically controversial issue. *Very* succinctly and simplifying, a neoclassical model can accommodate a negative labour supply elasticity only at the cost of an implausibly high income effect, and the early empirical findings of Camerer et al. [7] indicate precisely such a negative elasticity. This negative elasticity is informally explained with *income targeting*. Koszegi and Rabin's [33] (KR) use their reference dependence theory to explain this anomaly. However, Farber [21] rejects (on econometric and methodological grounds) the empirical finding, using a different dataset. In further work, Farber [22] formally introduces income targeting by drivers in a structural model, but on the basis of his data attributes low predictive power to the model. But in a recent breakthrough Crawford and Meng [12] (CM) devise an econometric specification of KR's theory with Farber's data, using income and leisure targeting and sample proxies for KR's rational expectations targets. This (assumed) target observability is the key to extract information from the data. In this way CM are able to explain non-neoclassical responses provided that (1) the gain-loss utility has a sufficiently large weight and (2) wage changes are *unanticipated* (while in the other cases the model predicts the textbook implication of nonnegative elasticity).

We apply our model to this situation as follows. We assume that the agent demands leisure always looking first at two types of properties: an 'income property' requiring income to be above a target level, and a 'leisure property' requiring leisure to be above a target level. These are very natural properties to look at, and as explained above the existence of 'target incomes' or 'target leisure levels' has been often considered in the relevant literature. The state determines the order in which the properties are looked at.

At first sight, it might appear that, because our model is lexicographic (unlike models with reference dependence and targeting, which have weights and trade-offs), the behaviour of the agent is entirely determined by the priority accorded to the two types of properties. For example one may suppose that agents giving priority to income react to drops in wages by *increasing* their labour supply, a negative elasticity that as explained is implausible in the textbook model. But in fact our model has a richer and more nuanced set of implications. The reason is simple. There is a qualitative difference between income properties and leisure properties. For *any* leisure property there always is a feasible leisure-income combination satisfying that property.

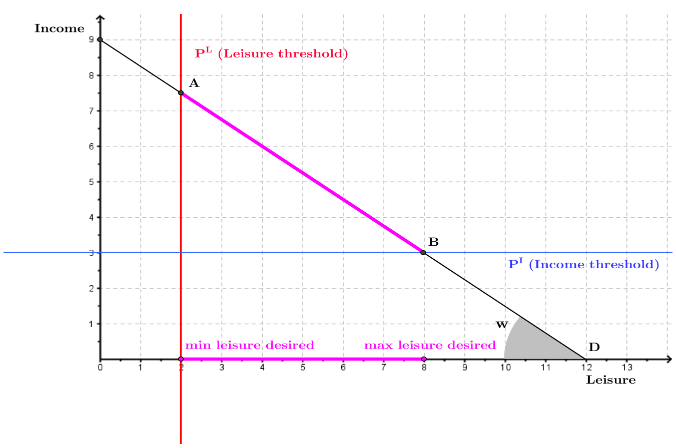

Figure 3: Labour supply

But, on the contrary, for *any* income property there is always a sufficiently low wage such that there is no leisure-income combination satisfying that property. This is a *general* observation, about target properties: one cannot (if he is sane) be 'unrealistic' with an *objective* constraint -like hours in a day-, but one can be unrealistic with a *subjective* constraint -like how much one thinks his work is worth. Because of this asymmetry, depending on the realisation of the wage even a state giving in principle priority to a target income will simply force the agent to move on to consider the next property, something that will never occur for leisure properties. Consider the situation in figure 3, depicting choices at some given wage w. The income and the leisure property, P^I and P^L, respectively, determine the same set AB of acceptable choices regardless of the order in which they are applied (we may imagine that a final selection within AB will be determined by other properties or by random factors, but we don't need to be specific in this short example). Suppose now that wage drops by a large amount to

$w' < w$, as illustrated in figure 4.

Once again, the state is irrelevant: the same set of acceptable alternatives CD is determined independently of the order, since if the income property P^I is applied first it will not have any bite. The set of acceptable choices does not contain any choice with lower leisure than the acceptable choices at the higher wage w, and contains many choices with higher leisure. This is broadly consistent with the effect predicted by the textbook model and with the empirical evidence analysed in Farber [21].

However, the state becomes crucial in determining the agent response to a milder drop in wage, to w'' with $w' < w'' < w$, as depicted in figure 5. Here it is still possible that the regular comparative statics result holds, if P^L is applied first. Yet if the agent is an 'income type' and instead applies P^I first, his response to the drop in wage is to *increase* labour supply to somewhere in the range corresponding to EF.

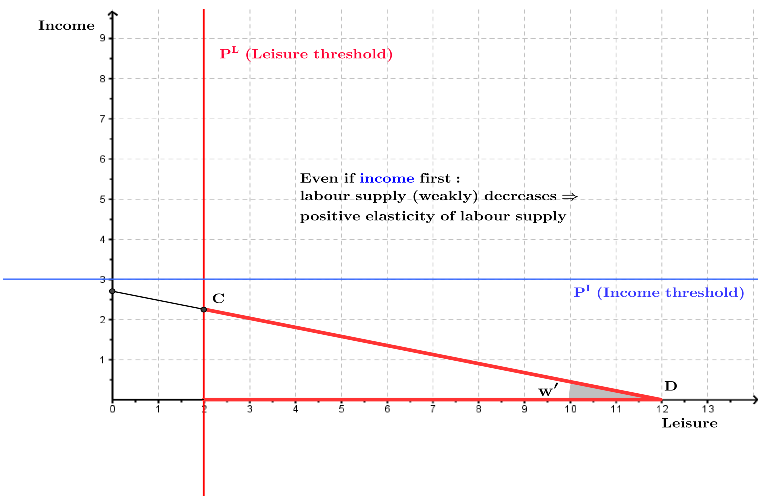

Figure 4: Labour supply and a large drop in wage

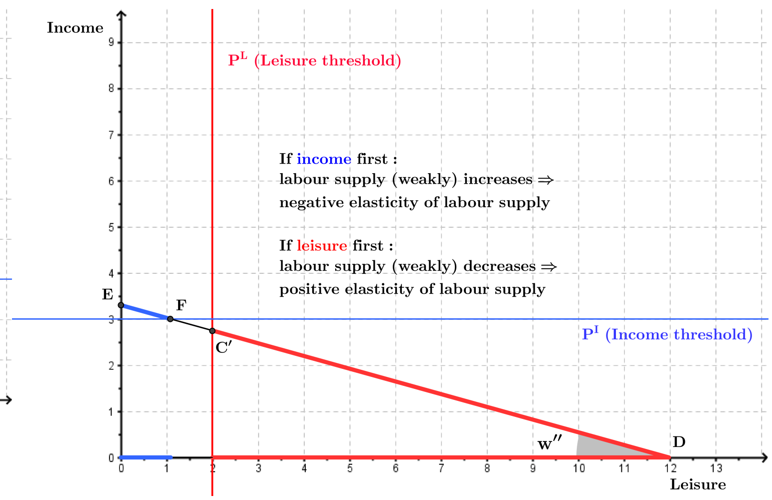

Figure 5: Labour supply with a modest drop in wage

Furthermore, regardless of whether P^L or P^I is applied first, our model predicts that labour supply will never range in the values corresponding to leisure levels identified by the FC′ range in figure 5 - any contrary observation would falsify our model, and the higher the leisure or wage threshold levels (i.e. the farther to the northeast the intersection between the two threshold properties moves), the larger the labour supply 'gap'.

Thus, the mechanism we have studied is consistent with textbook labour supply responses to large wage changes, but also allows a wider range of responses to milder wage changes, including negative wage elasticities. Negative wage elasticities are predicted when the agent targets income before leisure.

We note one interesting feature of our sketch model compared to the CM adaptation of KR's reference dependent model, which also accommodates both types of effects. In CM/KR, there is an element of rational expectations in that the reference targets are determined endogenously and 'correctly' by the agent, but the anomalous supply response can only occur for *unanticipated* wage changes. In our mechanism, on the contrary, *targets are exogenous unobserved variables* because they are very much like preferences in standard theory, that is fundamental primitives and not an additional variable in a utility framework. Thus, they are not necessarily rationally determined: but, on the other hand, the anomalous responses can occur for perfectly predicted wages without requiring expectation errors as in the reference dependence model.

In summary, then, our analysis points to two novel implications. First, the *size* of the wage change is a potentially important factor in the contrasting empirical evidence. Second, if state dependence explains labour supply then this supply can be 'behaviourally extreme', that is, for certain values of the wage there is a range of intermediate values of supply that are never chosen in any state: the worker either demands 'a lot' or 'little' labour, depending on the state. These hypotheses seem to merit further study both empirically and theoretically.

Of course, the above is just a sketch of a complete model, but it is precisely its simplicity coupled with the richness of implications that suggests its potential usefulness in more detailed frameworks. Similar applications of the theory can be made for example to saving decisions (along the lines of the example in section 3.1, explaining why visceral immediate consumption impulses may become more relevant at low rates of return on savings and why savers may be subject to behavioural swings); or to consumer theory (explaining the behaviour of a 'consumption-target driven' consumer).

6 Related literature

6.1 Multiself models

In view of Proposition 1, using a checklist that varies with A is equivalent to maximising a weak order that depends on A. So our model of menu driven state dependent choice may appear the same as that by Kalai et al. [29] (KRS), who assume that the decision maker maximises a preference relation (a weak order) which depends on the menu. But a moment's reflection shows that the two models are in fact distinct: after all, while our model is characterised by Togetherness, the KRS model can explain any choice observation.[16]

To understand the puzzle, since in the KRS model for any A we can simply pick a preference that puts the choice from A in the highest indifference class in A, it may happen that x is strictly preferred to y in A, but it is indifferent to y in another menu B when two different preferences are applied to A and B. Therefore it may happen that x is chosen and y is rejected from A, while both x and y are chosen from B, thus violating Togetherness. This highlights the centrality in our model of the *fixed nature of the mindset*. The fact that departures from 'rationality' in choice can only be attributed to state dependence, and not to the mindset, imposes, unlike the KRS model, some discipline on behaviour.

Similar observations hold for the recent works by Green and Hojman [24] and Ambrus and Rozen [3]. Green and Hojman describe an agent as a probability distribution over all possible preference orderings. Such preferences are then aggregated via a voting rule. If the voting rule satisfies a certain monotonicity property, this model can also explain any choice behaviour. Ambrus and Rozen study a very general model of a decision maker as a collection of utility functions ('selves'), which encompasses many other models in this vein. Each menu may activate a different aggregation procedure of the various selves. This aggregation procedure is constrained to satisfy some natural axioms, which however force the aggregation rule to incorporate some cardinal information: this contrasts with our and the other models mentioned in this section, which are purely ordinal. Ambrus and Rozen's central result is that with a sufficient number of selves any choice observation can be explained, in spite of the restrictions imposed on the aggregation procedure. This result suggests that multiselves models need to limit the number of allowable selves in order to exhibit observable restrictions. Replacing 'selves' with 'states', this intuition could also apply to our framework. Because our analysis specifies the components of a state, one obvious way to restrict a state is to limit the number of properties that constitute it (some environments may specify natural restrictions). We also have noted above how the assumption of nestedness of the states makes our model equivalent to utility maximisation.

[16] To be precise, KRS deal with choice functions. We are referring to the obvious extension of their ideas to choice correspondences, in which a weak order is maximised in each menu.

6.2 Choice with frames

As observed in section 4, the choice functions $\gamma(., <_m)$ when single-valued can be interpreted as choice with frames (Salant and Rubinstein [53]), where each state $<_m$ plays the role of a specific frame. Of course, in our framework a checklist is a very different object from a utility function, as a linear order of a fixed set of *properties* rather than of alternatives. However using Proposition 1 as a bridging result, we can associate with each $<_m$ a weak order on the set of alternatives. When the weak order is a strict linear order, such a substitution generates a choice function that Salant and Rubinstein term 'choice by salient consideration' (a single-valued choice function that maximises some frame dependent strict linear order on the alternatives). In the proof of Proposition 6 however we implicitly prove that a c with a environmental checklist can always be seen (in our domain) as the union of single-valued choice functions. In short, then, we can establish an observational equivalence between the choice correspondence induced by a choice by salient consideration as frames vary, and the choice correspondence c induced by $\gamma(., <_m)$ as states vary. The upshot is that, as a by-product, Proposition 6 also provides a characterisation (hitherto not available) of choice correspondences generated by salient consideration choice functions.

The fact remains, however, that it might not be possible to identify the set of frames needed to interpret an observed choice as framed with the states that generated it. For an extreme example, if x and y have the same properties and there is a *single* state, we still need at least *two* frames to generate, for example, the choice $c(\{x, y\}) = \{x, y\}$. This discrepancy would not go away even if we extended the frame model to allow for multivalued choices in each state, for reasons similar to those explaining the discrepancy between menu dependent utility maximisation and menu driven state dependent choice. For example, if x and y are as above, they will always appear together in any choice (since they must be 'indifferent' in any state), but one could instead define a frame in which x is ranked above y. So, at the conceptual level, the state dependence model remains distinct from the frame model. If states or frames could be observed, or if at least the choices corresponding to each state/frame could be observed, the two models could be easily told apart.

Finally, note that even the choice-observational correspondence between choice by salient considerations and environmental states is due to our assumption that properties are satisfied in a $0 - 1$ fashion (together with the bridging result of MMM [40]). For example, in a variation of our model in which each property can be satisfied in various degrees represented by a partial order (i.e. a 'state dependent' version of our

[42] model of choice by lexicographic semiorder) the correspondence would be lost.

6.3 Distinguishing state dependent from indecisive behaviour

Distinguishing between the *psychological reasons* that motivate behaviour is important in many economic contexts. In this section we compare state dependence with an important psychological driver, *indecisiveness* (Eliaz and Ok [18], Mandler [39]). Indecisiveness is conceived by Eliaz and Ok and Mandler as the lack of either a strict preference or an indifference.

A simple property characterises the choice behaviour of indecisive agents in Eliaz and Ok's [18] model:

WARNI (Weak Axiom of Revealed Non-Inferiority): If in a menu A there is an alternative x that is chosen, possibly from different menus, in the presence of each alternative in A, then x should be chosen from A. Formally, for all $A, B \in \Sigma$:

$[\forall y \in A \; \exists B \in \Sigma$ such that $x \in c(B), y \in B] \Rightarrow x \in c(A)$.

Eliaz and Ok assume that the agent maximises an incomplete preference relation, and thus interpret the fact that $x \in c(B)$ and $y \in B$ as revealing the non-inferiority of x compared to y (though not necessarily its superiority). This suits a situation in which an agent is undecided, rather than indifferent, between two alternatives (in which case he cannot rank them). It is easy to see that both our models of state dependent behaviour have different empirical predictions from Eliaz and Ok's model of indecisiveness. Call a c that satisfies WARNI but not WARP an *indecisive choice*.

Claim 2 *There are (multivalued) menu driven state dependent choices that are not indecisive,*[17] *and vice-versa. Similarly, there are (multivalued) environment driven state dependent choices that are not indecisive, and vice-versa.*

Other comparisons that it would be interesting to make are with the psychological phenomena of *inattention* (Masatlioglu et al. [46]), *categorisation* (Manzini and Mariotti [43]) and *rationalisation* (Cherepanov et al. [11]). However, all these theories, while axiomatised, are formulated for single-valued choices, whereas ours is crucially predicated on multiple observations of choice from menus. Therefore, pending a generalisation of those theories to choice correspondences, only the following basic claim can be made: there are menu driven state dependent choices that cannot be explained

[17] The claim is trivial for single-valued state dependent choices.

by inattention, categorisation or inattention, and not viceversa (menu driven state dependent choice is unrestrictive when single-valued) and there are choices driven by inattention, categorisation and rationalisation that are not environment driven state dependent choices, and not viceversa (sd-WARP reduces to WARP in the single-valued case).

7 Concluding remarks

In this paper we have proposed a versatile framework to study the dependence of choices on psychological states. The language of properties we have adopted allows one to talk about such states in a more nuanced way, and with more direct psychological interpretations, than the language of utility. The framework can accommodate much non-standard behaviour, yet we have shown that it is not empirically vacuous.

'Psychological' theories of choice that use primitives different from a standard utility function face a potential problem: namely, that it's hard to tell precisely how different they are from the textbook model.[18] While this problem may be legitimately addressed in various ways, we have found it useful to address it through an *axiomatic* characterisation. In this way we have (a) pinned down the precise regularities in choice behaviour implied by the state swings of our framework - generating testable implications on choice data - and (b) located our models in the 'logical space' of standard revealed preference axioms, facilitating comparison both with the textbook model and with other psychological theories of choice. Our models are, in a sense made precise, 'between' the standard model and certain weakened versions of it. In addition, the models are tractable: with appropriate restrictions, a rich set of specific economic implications can be derived.

In both models we have considered, the idea at the heart of their testability and distinctiveness is that an observer can garner data on multiple choices from the same menu (with single-valued observations, our first model would be empirically vacuous and our second model would reduce to the standard utility maximisation model). These multiple observations of choice from a menu can be taken literally, as for a consumer observed to pick different items from a shelf in different shopping days, taxi drivers choosing different hours of work in different days at the same daily wage, or experimental subjects approving multiple options. We hope that our results will spur

[18] See e.g. Gul and Pesendorfer [25] for a surprising application of the revealed preference method to a psychological model.

the collection of this type of data. But multiple observations of choice can also be interpreted as deriving from a (possibly estimated) *stochastic* choice function. In this case, the choosable alternatives in $c(A)$ can be interpreted for example as all those alternatives that are chosen with positive probability, or as those that are chosen with maximum probability, or any of the intermediate possibilities. Then, depending on the specific connection made between deterministic and stochastic choice function, our characterisation results for the deterministic model will translate into different restrictions on the probabilistic one.[19] Obviously, the data needed to test our theory are less demanding than the collection of stochastic choice data - we just need to tell *whether* an alternative is choosable, without worrying about the precise frequency -, yet even the latter can be collected without too much difficulty (see e.g. Caplin and Dean [9] for a recent design that even elicits state-dependent stochastic choices).

In practice, a state is not entirely unpredictable: psychological research may help to identify correlations between environmental and personal variables with states, and states with choice, so that additional elements of predictability in choice can be identified. For example, in the taxi driver application of section 5.1, one might conjecture a relationship between weather and psychological state, so that weather could be used as an observable proxy for states (or, conversely, one could use choice data to *infer* a connection between choice and state, conditional on the theory being valid). Our contribution has been to identify what can be predicted *exclusively* in terms of an economic choice model.

While a psychological state affects choice, undoubtedly choice affects the psychological state, too. There is a subtle two-way interaction between psychological states and choices. At present it is not clear how this interaction can be modelled, yet some progress has been made.[20] Our paper is a first step towards the formal modelling of state dependent choice behaviour.

[19] A classical technical treatment of the relationship between deterministic choice correspondences and stochastic choice functions is Fishburn [23].

[20] See for example, Dalton and Ghosal [13], who resolve the interaction through an elegant equilibrium analysis, and Dillenberger and Rozen [16], where disappointment and elation are moods that affect risk aversion, and are determined endogenously based on how risk unfolds over time (generating a history of disappointment and elation states).

References

[1] Aizerman, M.A. and Fuad Aleskerov (1995), *Theory of choice*, North Holland, Amsterdam and New York.

[2] Aizerman, M.A. and A.V. Malishevski (1981) “General theory of best variants choice: some aspects”, *IEEE Transactions on Automatic Control*, 26:1030–40.

[3] Ambrus, A. and K. Rozen (2013) “Rationalizing Choice with Multi-Self Models”, forthcoming in the *Economic Journal*.

[4] Apesteguia, J. and M. A. Ballester (2013) “Choice by Sequential Procedures”, *Games and Economic Behavior*, 77 (1): 90–99.

[5] Bernheim, B. D. and A. Rangel (2009) “Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics”, *Quarterly Journal of Economics*, 124:51-104.

[6] Berry, S., J. Levinsohn and A. Pakes (1995) “Automobile Prices in Market Equilibrium”, *Econometrica* 63(4): 841-90.

[7] Camerer, C., L. Babcock, G. Loewenstein and R. Thaler (1997) “Labor Supply of New York City Cabdrivers: One Day at a Time”, *Quarterly Journal of Economics*, 112: 407–41.

[8] Caplin, A. and J. Leahy (2001) “Psychological expected utility theory and anticipatory feelings”, *Quarterly Journal of Economics*, 116: 55-79.

[9] Caplin, A and M. Dean (2013) “Rational Inattention and State Dependent Stochastic Choice”, mimeo, NYU.

[10] Capra, C. Monica (2004) “Mood-driven behavior in strategic Interactions”, *American Economic Review Paper and Proceedings* 94(2): 367-372.

[11] Cherepanov, V., T. Feddersen and A. Sandroni (2013) “Rationalization”, *Theoretical Economics*, 8 (3): 775–800.

[12] Crawford, V. P. and J. Meng (2011) “New York City cab drivers’ labor supply revisited: reference-dependent preferences with rational-expectations targets for hours and income”, *American Economic Review* 101: 1912–1932.

[13] Dalton, P. and S. Ghosal (2008) “Behavioral decisions and welfare”, Warwick Economic Research papers n. 834, University of Warwick.

[14] Dekel, E., B. Lipman and A. Rustichini (2001) ”Representing Preferences with a Unique Subjective State Space,” *Econometrica* 69: 891-934.

[15] Dietrich, F. and C. List (2013) “Reason Based Rationalization”, mimeo, London School of Economics.

[16] Dillenberger, D. and K. Rozen (2013) “History-Dependent Risk Attitude”, working paper, University of Pennsylvania,

http://economics.sas.upenn.edu/˜ddill/DillenbergerRozen_HDRA%20FEB2013.pdf.

[17] Ehlers, L. and Y. Sprumont (2008) "Weakened WARP and top-cycle choice rules", *Journal of Mathematical Economics* 44 (1) 87-94.

[18] Eliaz, K. and E. Ok. (2006) "Indifference or indecisiveness? Choice theoretic foundations of incomplete preferences", *Games and Economic Behavior* 56: 61–86.

[19] Erber, R., M. Wang and J. Poe (2004). Mood regulation and decision making: Is irrational exuberance really a problem? In I. Brocas and J.D. Carrillo (eds.), '*The psychology of economic decisions*' (II:197-210), Oxford University Press.

[20] Ergin, H. and T. Sarver (2010) "A Unique Costly Contemplation Representation", *Econometrica* 78: 1285-1339.

[21] Farber, H.S. (2005) "Is tomorrow another day? The labor supply of New York City cabdrivers", *Journal of Political Economy*, 113: 46-82.

[22] Farber, H.S. (2008) "Reference-dependent preferences and labor supply: the case of New York City taxi drivers", *American Economic Review*, 98: 1069–82.

[23] Fishburn, P.J. (1978) "Choice probabilities and choice functions", *Journal of Mathematical Psychology* 18: 205-19.

[24] Green, J. and D. Hojman (2008) "Choice, rationality and welfare measurement", mimeo, Harvard IEResearch Discussion Paper n.2144.

[25] Gul, F. and W. Pesendorfer (2005) "The Revealed Preference Implications of Reference Dependent Preferences", mimeo, Princeton University.

[26] Ifcher, J. and H.S. Zarghamee (2011) "Happiness and time preference: the effect of positive affect in a random-assignment experiment", *American Economic Review*, 101(7): 3109-29.

[27] Isen, A.M. (2000) 'Positive affect and decision making', In M. Lewis & J.M. Havieland (eds.), *Handbook of emotions* (2:417-35). London: Guilford.

[28] Kahn, B.E., and A.M. Isen (1993) 'The influence of positive affect on variety seeking among safe, enjoyable products', *Journal of Consumer Research*, 20:257-70.

[29] Kalai, G., A. Rubinstein and R. Spiegler (2002) "Rationalizing choice functions by multiple rationales" *Econometrica*, 70: 2481-2488.

[30] Kirchsteiger, G., L. Rigotti and A. Rustichini (2006), "Your morals might be your moods", *Journal of Economic Behavior & Organization*, 59:155-72.

[31] Kliger, D. and Levy, O. (2003) "Mood-induced variation in risk preferences", *Journal of Economic Behavior & Organization*, 52:573-84.

[32] Köszegi, B. (2010) "Utility from anticipation and personal equilibrium", *Economic Theory*, 44:415-44.

[33] Köszegi, B. and M. Rabin (2006) "A model of reference-dependent preferences", *Quarterly Journal of Economics* 121: 1133-65.

[34] Kreps, D. M. (1979) "A Representation Theorem for 'Preference for Flexibility' ", *Econometrica*, 47: 565-577.

[35] Laibson, D. (2001) "A cue-theory of consumption", *Quarterly Journal of Economics*, 116:81-119.

[36] Lancaster, K. J. (1966) "A new approach to consumer theory ", *The Journal of Political Economy*, 74 (2): 132-157.

[37] Litvakov, B.M. (1981) "Minimal representation of joint-extremal choice of options", Automation and Remote Control 1: 182-184.

[38] Mayer, J., Y. Gaschke, D. Braverman and T. Evans (1992) "Mood-congruent judgment is a general effect", *Journal of Personality and Social Psychology*, 63: 119-132.

[39] Mandler, M. (2009) "Indifference and Incompleteness distinguished by rational trade", *Games and Economic Behavior* 67:300-14.

[40] Mandler, M., P. Manzini and M. Mariotti (2012) "A million answers to twenty questions: choosing by checklist", *Journal of Economic Theory*, 147:71–92.

[41] Manzini, P. and M. Mariotti (2007) "Sequentially rationalizable choice", *American Economic Review*, 97:1824-39.

[42] Manzini, P. and M. Mariotti (2012) "Choice by lexicographic semiorders", *Theoretical Economics*, 7: 1-23.

[43] Manzini, P. and M. Mariotti (2012) "Categorize then choose: boundedly rational choice and welfare", *Journal of the European Economics Association*, 10: 1141–65.

[44] Manzini, P. and M. Mariotti (2013) "Stochastic choice and consideration sets", *Econometrica*, 89(3): 1153-1176.

[45] Manzini, P., M. Mariotti and C.J. Tyson (2012) "Two-stage threshold representations", *Theoretical Economics*, 8 (3): 875–882.

[46] Masatlioglu, Y., D. Nakajima and E. Ozbay (2012) "Revealed attention", *American Economic Review* 102:2183–205.

[47] Mittal, V., and W.T. Ross (1998) "The impact of positive and negative affect and issue framing on issue interpretation and risk taking", *Organizational Behavior and Human Decision Processes*, 76: 298-324.

[48] Nygren, T.E. (1998) "Reacting to perceived high- and low-risk win-lose opportunities in a risky decision-making task: is it framing or affect or both?", *Motivation and Emotion*, 22:73-98.

[49] Nygren, T.E., A.M. Isen, P.J. Taylor, and J. Dulin (1996) "The influence of positive affect on the decision rule in risk situations: Focus on outcome (and especially avoidance of loss) rather than probability", *Organizational Behavior and Human Decision Processes*, 66: 59-72.

[50] Oswald, A.J., E. Proto and D. Sgroi (2009) "Happiness and productivity", IZA DP

n. 4645.

[51] Rick, S. and G. Loewenstein (2008) "The role of emotion in economic behavior", ch. 9 in *The Handbook of Emotion*, 3^{rd} ed, M. Lewis, J. Haviland-Jones and L. Feldman-Barrett (Eds.), New York, Guilford.

[52] Rubinstein, A. and Y. Salant (2012) "Eliciting welfare preferences from behavioral datasets", *Review of Economic Studies*, 79 (1): 375-387.

[53] Salant, Y. and A. Rubinstein (2008) "(A, f):Choice with frames", *Review of Economic Studies*, 75: 1287–96.

[54] Thayer, R.E. (2000) 'Mood'. *Encyclopedia of Psychology*, Washington, D.C.: Oxford University Press and American Psychological Association.

[55] Thayer, R.E. (2001) 'Calm energy: how people regulate mood with food and exercise', New York, Oxford University Press.

[56] Williams, S., and Y. W. W. Voon (1999) "The effects of mood on managerial risk perceptions: Exploring affect and the dimensions of risk", *Journal of Social Psychology*, 139: 268-287.

Appendices

A Proofs

Proof of Proposition 2. We show the following: let c and d be two choice functions on Σ that have, respectively the sd-cheklists $\left(\Gamma, \{<_A\}_{A\in\Sigma}\right)$ and $\left(\Gamma, \{<'_A\}_{A\in\Sigma}\right)$; then $c = d$.

Suppose that $c\,(A) \neq d\,(A)$ for some $A \in \Sigma$ and in particular let (possibly relabeling the choice functions) $x \in c\,(A)$ and $x \notin d\,(A)$. The latter implies that there exist $y \in A$ and $P \in \Gamma$ such that $x \notin P$ and $y \in P$. For $x \in c\,(A)$ it must then be the case that there exists $Q <_A P$ and $z \in A$ such that $y \notin Q$ and $z \in Q$. If $P \subset Q$ this is incompatible with $y \in P$, and if $Q \subset P$ then $x \notin Q$. Therefore $x \notin S_A(Q, <_A)$ and $x \notin c\,(A)$, a contradiction.

So any sequence of the properties in the mindset Γ generates the same behaviour and by Proposition 1 the behaviour generated by any particular sequence maximises a weak order, as claimed. ∎

Proof of Proposition 3. Suppose that c satisfies Togetherness. Define a relation $\approx$ on X by $x \approx y$ iff there is no $A \in \Sigma$ such that $x \in c\,(A)$ and $y \in A \backslash c\,(A)$ or $y \in c\,(A)$ and $x \in A \backslash c\,(A)$. The relation $\approx$ is obviously reflexive and symmetric. To see that it is also transitive, suppose that $x \approx y \approx z$ and that $x \in c\,(A)$ and $z \in A$ for some $A \in \Sigma$. We show that $z \in c\,(A)$.

Since $x \approx y$ we have $x \in c\,(\{x,y,z\})$ if and only if $y \in c\,(\{x,y,z\})$ ($\{x,y,z\}$ is in the domain by assumption), and similarly $y \approx z$ implies that $y \in c\,(\{x,y,z\})$ if and only if $z \in c\,(\{x,y,z\})$. Therefore if $x \in c\,(\{x,y,z\})$ then $c\,(\{x,y,z\}) = \{x,y,z\}$. Therefore by Togetherness $x \in c\,(A)$ and $z \in A$ imply $z \in A$. If instead $x \notin c\,(\{x,y,z\})$, then $c\,(\{x,y,z\}) = \varnothing$, a contradiction. $\approx$ is therefore an equivalence relation and it partitions the set of alternatives into equivalence classes, which we denote $[x] = \{y \in X : y \approx x\}$.

Given $A \in \Sigma$, take any $x \in c\,(A)$ and let $P_A = [x]$. Note that P_A is uniquely defined, and let the mindset be $\Gamma = \{P_A : A \in \Sigma\}$. Since $\approx$ is an equivalence, we have $P_A \cap P_B = \varnothing$ for all distinct menus $A, B \in \Sigma$. Let the state $<_A$ be any linear order for which $P_A <_A P$ for all $P \in \Gamma \backslash P_A$. Then $A \cap P_A = c\,(A)$ (for any $y \notin c\,(A)$ and $x \in P_A$ it cannot be $y \approx x$ by the definitions of $\approx$ and P_A). And since for all $P \in \Gamma \backslash P_A$ we have $P_A \cap P = \varnothing$, it follows that, for all $P \in \Gamma$, $S_A\,(P, <_A) = c\,(A)$.

Conversely, let $\{\Gamma, \{<_A\}_{A\in\Sigma}\}$ be a menu checklist for c. Suppose $x, y \in c(A)$ for some $A \in \Sigma$, and suppose by contradiction that, for some $B \in \Sigma$, $y \in c(B)$ and $x \in B \backslash c(B)$. Then there exists $P \in \Gamma$ such that $y \in P$ and $x \notin P$. By definition of having a checklist there exists $Q \in \Gamma$ such that $S_A(Q) = S_A(R)$ for all $R \in \Gamma$ with $Q <_A R$. This cannot be true if $Q <_A P$. On the other hand, if $P <_A Q$ it cannot be $x, y \in c(A)$, a contradiction. ■

Proof of Proposition 4: If $x \in \gamma(B, <_m)$ then x must be in any $S_B(P, <_m)$, and therefore in any $S_A(P, <_m)$, so that $x \in \gamma(A, <_m)$ whenever $x \in A$: so $\gamma(B, <_m) \cap A \subseteq \gamma(A, <_m)$. And if $x \notin \gamma(B, <_m)$ and $\gamma(B, <_m) \cap A \neq \varnothing$, then there exists $y \in \gamma(B, <_m) \cap A$ that has a property which x does not have. So there is P for which $x \notin S_A(P, <_m)$, and consequently $x \notin \gamma(A, <_m)$. This shows that $\gamma(A, <_m) \subseteq \gamma(B, <_m) \cap A$, and we conclude that $\gamma(A, <_m) = \gamma(B, <_m) \cap A$. ■

Proof of Claim 1 Let $\Gamma = \{P, Q, R\}$, $M = \{<_m, <_n\}$, $P <_m Q <_m R$ and $R <_n Q <_n P$. Suppose $P = \{x, z\}$, $Q = \{y, z\}$ and $R = \{x, y\}$. Then

$$\begin{array}{ll} \gamma(\{x,y\}, <_m) = \{x\} & \gamma(\{x,y\}, <_n) = \{y\} \\ \gamma(\{x,z\}, <_m) = \{z\} & \gamma(\{x,z\}, <_n) = \{x\} \end{array}$$

and therefore

$$\begin{array}{lcl} c(\{x,y\}) = \{x,y\} & \multirow{2}{*}{\text{But}} & \gamma(\{x,y,z\}, <_m) = \{z\} \\ c(\{x,z\}) = \{x,z\} & & \gamma(\{x,y,z\}, <_n) = \{y\} \end{array}$$

so that $c(\{x, y, z\}) = \{y, z\}$. We conclude that c violates both Expansion and Property β.

Proof of Proposition 6. We will find it convenient to write sd-WARP in an equivalent way:[21]

sd-WARP (restated): $c(A) \subseteq B \Rightarrow c(B) \cap A \subseteq c(A)$.[22]

Necessity. As a preliminary, we say that 'x $m-$tops y', written $xT_m y$, if there is a

[21]To see this, let sd-WARP hold, and suppose that $C(A) \subset B$ but that in contradiction there exists some x such that $x \in (C(B) \cap A) \backslash C(A)$. Since $C(A) \subset B$ and $x \in A \backslash C(A)$, sd-WARP requires $x \notin C(B)$, contradiction. For the other direction, let sd-WARP (restated) hold, and suppose that $x \in A \backslash C(A)$, $C(A) \subset B$ but that in contradiction $x \in C(B)$. Then $x \in (C(B) \cap A) \backslash C(A)$, an immediate contradiction of sd-WARP (restated).

[22]A small choice theoretic observation: this formulation makes it clear that sd-WARP is a stronger version of the classic axiom by Aizerman (see Aizerman and Malishevski [2]), which adds to the premise in sd-WARP the requirement that the sets A and B are nested, i.e. $B \subset A$.

state $<_m$ and a property P_i such that $x \in P_i$, $y \notin P_i$ and $y \in P_j \Rightarrow x \in P_j$ for all P_j such that $P_j <_m P_i$. Observe that for all $D \in \Sigma$ and $<_m \in M$, $x \in \gamma(D, <_m)$ only if there is no $y \in D$ such that $yT_m x$.

Let $A, B \in \Sigma$ be such that $c(A) \subseteq B$. The statement of the proposition is trivially true if $A \cap \gamma(B, <_m) = \varnothing$ for all $<_m \in M$ (in which case $A \cap c(B) = \varnothing$), so suppose that $A \cap \gamma(B, <_m) \neq \varnothing$ for some $<_m \in M$. Then

$$\begin{aligned} A \cap c(B) &= A \cap \bigcup_{<_m \in M} \gamma(B, <_m) \\ &= \bigcup_{<_m \in M} A \cap \gamma(B, <_m) \subseteq \bigcup_{<_m \in M} \gamma(A, <_m) \\ &= c(A) \end{aligned}$$

where the inclusion is proved with the following reasoning. Since $c(A) \subseteq B$, for all $<_m \in M$ we have $\gamma(A, <_m) \subseteq B$. So in particular there is no $y \in A \backslash B$ that $m-$tops any $x \in \gamma(A, <_m)$. Therefore for all $x \in \gamma(B, <_m) \cap A$ we also have $x \in \gamma(A, <_m)$ (if not, there would exist $y \in A \backslash B$ with $yT_m x$). We conclude that, for all $<_m \in M$, $\gamma(B, <_m) \cap A \subseteq \gamma(A, <_m)$, from which the desired inclusion follows.

Sufficiency. Let sd-WARP hold. We construct an environmental checklist explicitly, then show that it retrieves $c(A)$ for each menu $A \in \Sigma$. Let $\Gamma = \left\{ \{x\}_{x \in X} \right\}$, and let $|X| = n$.

An $a-$path is a sequence $a = \{x_i\}_{i=1,...n}$ of distinct alternatives $x_1, x_2, ...x_n$ defined recursively as follows. $x_1 \in c(X)$ and, for all $i > 1$, $x_i \in c(X \backslash \{x_1, ..., x_{i-1}\})$. Denote by α the collection of $a-$paths, and note that each $a-$path covers all of the alternatives in X . Construct M by setting, for each $a \in \alpha$:

$$\{x_i\} <_a \{x_j\} \text{ if and only if } i < j \text{ and } x_i, x_j \in a$$

We now show that this construction retrieves choice.

Fix an arbitrary menu $A \in \Sigma$, let $x \in c(A)$, and suppose by contradiction that for each $<_m \in M$, there is an alternative w such that $\{w\} <_m \{x\}$. For each $<_m$ let $\{y_m\}$ denote the $<_m$ $-$maximal property in A, that is $\{y_m\} <_m \{z\}$ for all $z \in A \backslash \{y_m\}$. By construction we have that $y_m \in \gamma(A_m, <_m)$ where $A_m = \{y_m\} \cup \{z \in X : \{y_m\} <_m \{z\}\}$. Observe that by assumption there is no $<_m$ such that $x \in \gamma(A_m, <_m)$ (otherwise $\{x\}$ would be maximal in A for some state). Moreover, by construction it must also be that $A \subseteq A_m$, for otherwise it would not be true that $\{y_m\} <_m \{z\}$ for all $z \in A \backslash \{y_m\}$.

If for any of the A_m it is the case that $c(A_m) \subseteq A$, then by sd-WARP it would follow that $c(A) \cap A_m \subseteq c(A_m)$, contradicting $x \notin \gamma(A_m, <_m)$. So suppose not, so that $c(A_m) \backslash A \neq \varnothing$, and consider $A_{m1} = A_m \backslash \{z_1\}$ where $z_1 \in c(A_m) \backslash A$. As before, either $c(A_{m1}) \subseteq A$, so that the contradiction $c(A) \cap A_{m1} \subseteq c(A_{m1})$ follows; or $c(A_{m1}) \backslash A \neq \varnothing$. More in general, proceed recursively setting $A_{mj} = A_{mj-1} \backslash \{z_j\}$ where $z_j \in c(A_{mj}) \backslash A$ whenever $c(A_{mj}) \backslash A \neq \varnothing$ and $j > 1$. At each step either $c(A_{mj}) \subseteq A$, implying $c(A) \cap A_{mj} \subseteq c(A_{mj})$; or $c(A_{mj}) \backslash A \neq \varnothing$. Since X is finite there exists a j^* such that $c(A_{mj^*}) \backslash A = \varnothing$, generating the desired contradiction.

Suppose now that $x \in A \backslash c(A)$, and that in contradiction there exists some state $<_m \in M$ such that $x \in \gamma(A, <_m)$. By construction it must be that $x \in c(B)$ where $B = \{x\} \cup \{y \in X : \{x\} <_m \{y\}\}$, and that $A \subseteq B$. If $A = B$ we have an immediate contradiction. Otherwise, then $c(A) \subset B$, so that by sd-WARP $c(B) \cap A \subseteq c(A)$ also follows, implying $x \in c(A)$, a contradiction. ■

Proof of Claim 2. By example:

Example 4 *(Menu driven state dependent but not indecisive)* $c(\{x,y,z\}) = \{x,y\}$, $c(\{x,y\}) = \{x,y\}$, $c(\{x,z\}) = \{x\}$, $c(\{y,z\}) = \{z\}$.

Example 5 *(Environment driven state dependent but not indecisive)* $c(\{x,y,z\}) = \{x,y\}$, $c(\{x,y\}) = \{x,y\}$, $c(\{x,z\}) = \{x,z\}$, $c(\{y,z\}) = \{y,z\}$.

Example 6 *(Indecisive but neither menu driven state dependent nor environment driven state dependent)* $c(\{x,y,z\}) = \{y\}$, $c(\{x,y\}) = \{x,y\}$, $c(\{x,z\}) = \{z\}$, $c(\{y,z\}) = \{y\}$.

B Examples

B.1 Property α is not sufficient for environment driven state dependent choice

Suppose $c(\{x,y,z\}) = \{x\}$ and $c(\{x,y,w\}) = \{x,y\}$. Although these two choices do not violate Property α, it is not possible to find a mindset Γ, a set of states M and a choice function γ such that

$$c(A) = \bigcup_{<_m \in M} \gamma(A, <_m)$$

To see this, suppose to the contrary that $y \in \gamma(\{x, y, w\}, <_m)$ for some $m \in M$. Since $y \notin c(\{x, y, z\})$, it must be that there is an alternative $i \in \{x, y, z\}$ such that $i \in P_i$, $y \notin P_i$ and $y \in P_j \Rightarrow i \in P_j$ for all P_j such that $P_j <_m P_i$. If $i = x$, then it could not be that $y \in \gamma(\{x, y, w\}, <_m)$; while if $i = z$, then either $x \notin \gamma(\{x, y, w\}, <_m)$, or it must be that $z \in c(\{x, y, z\})$. In either case we have a contradiction.

B.2 sd-WARP implies Property α

Let sd-WARP hold, and suppose that there are sets A and B such that $A \subset B$ but that in contradiction to Property α there is some $x \in C(B) \cap A$ such that $x \notin C(A)$. Since $A \subset B$ it also follows that $C(A) \subset B$, which together with $x \in A \backslash C(A)$ and sd-WARP implies $x \notin c(B)$, contradiction.

B.3 Independence of the two state dependent choice models

The models we have considered in the paper (menu and environment driven states) are logically independent. We illustrate this with a simple example.[23] Let $A = \{h_s, h_L, u_s\}$ and $B = A \cup \{u_L\}$, where h_i and u_i stand for healthy and unhealthy food items, with the index 1 denoting a smaller portion than index 2. Let $c_1(A) = c_2(A) = \{h_s, h_L\}$, $c_1(B) = \{u_s, u_L\}$ and $c_2(B) = \{h_L, u_L\}$. Then c_s is menu driven but not environment driven state dependent (Togetherness holds and sd-WARP fails), while c_2 is environment driven but not menu driven state dependent (sd-WARP holds and Togetherness fails). The states for c_s could be e.g. $\{h_s, h_L\} <_A \{u_s, u_L\}$ and $\{u_s, u_L\} <_B \{h_s, h_L\}$, in line with the explanation that the decision maker can stick to healthy food when they are the majority, but adding an extra unhealthy food item switches the state to 'gluttony'. And for c_2 we could have $\{h_s, h_L\} <_c \{h_L, u_L\} <_c \{u_L\}$ when the decision maker is in a cool state, and sticks to his diet; while in a depressed state he seeks satisfaction in large portion sizes, e.g. $\{u_L\} <_d \{h_s, h_L\} <_d \{h_L, u_L\}$.

[23] This example is loosely based upon B. Wansink (1994) 'Antecedents and mediators of eating Bouts', *Family and Consumer Sciences Research Journal*, 23(2): 166–182, who studies the determinants of 'eating bouts'. An eating bout is a splurge on food that is some multiple (three times the usual amount in this study). Both external cues (i.e. menu composition) and internal states (i.e. moods) are invoked as alternative triggers for such eating bouts.

C An alternative characterisation of menu driven state dependent choice

A slightly different angle on the behavioural restriction of menu driven state dependent choice is the following:

All or Nothing: If the choices from two different menus overlap, then the choice from one menu consists of those available alternatives that are chosen from the other menu. Formally, for all $A, B \in \Sigma$: $c(A) \cap c(B) \neq \varnothing \Rightarrow c(A) \cap c(B) = c(A) \cap B$.

All or Nothing describes either a form of 'behavioural discontinuity' or of 'behavioural inertia', excluding other possibilities. That is, when moving from a menu A to a different menu B, either the agent's behaviour changes abruptly (no alternative is chosen from both menus) or whatever was originally chosen in A and is still available in B, it remains chosen, and no new alternatives are added to the choice.[24]

Corollary 1 *A choice function is a menu driven state dependent choice if and only if it satisfies All or Nothing.*

Proof. We show the equivalence of Togetherness and All or Nothing. Suppose that Togetherness holds and that $c(A) \cap c(B) \neq \varnothing$. Obviously for any $x \in c(A) \cap c(B)$ we have $x \in c(A) \cap B$, that is $c(A) \cap c(B) \subseteq c(A) \cap B$. For the converse inclusion, for any $x \in c(A) \cap B$ either $c(A) \cap c(B) = \{x\}$ or there exists $y \neq x$ with $y \in c(A) \cap c(B)$ and so by Togetherness $x \in c(B)$ (otherwise, $x \in B \backslash c(B)$ would violate Togetherness). This shows that $c(A) \cap B \subseteq c(A) \cap c(B)$ and we conclude that $c(A) \cap c(B) = c(A) \cap B$.

Conversely, suppose that Togetherness is violated, that is there exist $A, B \in \Sigma$ and $x \in A \backslash c(A)$, $y \in c(A)$, $y \in c(B)$ but $x \in c(B)$. Then $c(A) \cap c(B) \neq \varnothing$. Moreover, $x \notin c(A) \cap c(B)$ while $x \in c(B) \cap A$, so that $c(A) \cap c(B) \neq c(B) \cap A$, violating All or Nothing. ■

[24]Contrast again the restriction imposed by All or Nothing with that imposed by WARP, which can be written as $A \cap c(B) \neq \varnothing \Rightarrow c(A) \cap c(B) = c(A) \cap B$ (the same conclusion of All or Nothing from a weaker premise).

D A technical remark on the relation of Proposition 6 with a theorem by Litvakov

We establish a connection with a not very well-known result by Litvakov [37], asserting that if a choice function c satisfies Property α and Chernoff's Postulate 5 (if $c(B) \subseteq A \subseteq B$ then $c(A) = c(B)$),then it can be expressed as the union of choice correspondences c_i (i.e. $c(A) = \cup_i c(A)$ for all A), each of which satisfies IIA. Because Litvakov [37] is written in Russian, a more useful general reference is Aizerman and Aleskerov [1], Theorem 5.6(b) (and Theorem 5.2.1 for the other direction).

The following result can be proved (details available from the authors):

Claim 3 *A choice function c satisfies sd-WARP if and only if it satisfies Property $\alpha$and Chernoff's Postulate 5.*

Then, it can be seen that Proposition **6** can also be derived indirectly as the combined consequence of Litvakov's theorem, the above claim, and Proposition 1 in this paper.